DOCUMENT RESUME

ED 101 488 EC 070 635

AUTHOR Keating, Daniel P.
TITLE Precocious Cognitive Development at the Level of Formal Operations.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE [Oct 73]
NOTE 93p.; Ph.D. Dissertation, The Johns Hopkins University

EDRS PRICE MF-$0.76 HC-$4.43 PLUS POSTAGE
DESCRIPTORS *Comparative Analysis; *Developmental Psychology; Exceptional Child Research; *Gifted; *Intelligence Tests; Intermediate Grades; Junior High Schools; Students; *Task Performance
IDENTIFIERS Piaget (Jean)

ABSTRACT

One hundred nine fifth and seventh graders, classified as either bright or average, were tested to determine the relationship of intelligence (as defined by scores on psychometrically derived tasks) and developmental precocity (defined in terms of Piagetian theory). Specific measures used were the Iowa Tests of Basic Skills, Raven's Standard Progressive Matrices, and Piagetian tasks (conservation of volume, displacement, the balance, and the period of a pendulum). The major finding was that students who scored higher on psychometric measures of intelligence were also developmentally advanced in Piaget's sequence of cognitive developmental stages. The finding did not contradict Piagetian theory since out-of-sequence successes were not observed. (Author/GW)

ED101488

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

PRECOCIOUS COGNITIVE DEVELOPMENT AT THE LEVEL OF FORMAL OPERATIONS

By

Daniel P. Keating

A dissertation submitted to the Johns Hopkins University in conformance with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

1973

EC076685

ABSTRACT

Theorizing about the nature of intelligence has a long and varied history, but theories which have been based on systematic observation of some sort are relatively recent. Two traditions currently exert a major influence on research and speculation about intelligence within psychology: the psychometric tradition and the Piagetian or Genevan tradition. They are clearly differentiated in their orientation, the theories they generate, and the methods which they employ.

This research examines the relationship of "brightness" or "intelligence" as defined by scores on psychometrically derived tasks and developmental precocity in terms of the stage theory of Piaget. The age group for this study is early adolescence, specifically fifth and seventh graders. The specific psychometric measures are the Iowa Tests of Basic Skills and Raven's Standard Progressive Matrices. The Piagetian tasks are conservation of volume (advanced concrete level task), displacement, the balance, and the period of a pendulum (all formal operational tasks).

The major finding was that students who scored higher on psychometric measures of intelligence were also developmentally advanced in Piaget's sequence of cognitive developmental stages. This does not contradict the theory of Piaget, since out of sequence successes were not observed. The investigation was an

attempt to initiate research which examined the relationship of differential and generalizing perspectives on human intelligence.

PREFACE

It is difficult to trace the evolution of the formulation of a research problem, and to acknowledge adequately the intellectual debts incurred in its development. Thus the following can hardly be considered an all-inclusive list. Most importantly, I wish to thank Professor Julian Stanley, not only for his welcome input into this research, but also for his tremendous support and encouragement generally. Professor Rober Webb contributed generously in the development of the research reported here, and stimulated my interest in the theoretical aspects of intellectual precocity. Professors Robert Hogan and John Guthrie offered helpful comments at an early stage of this research.

The National Science Foundation supported the research reported here under its program for Doctoral Dissertation Research in Psychology. Many individuals were helpful at various stages in the research, and their contributions are acknowledged at appropriate points in the manuscript, but I wish to mention especially Rosalind Schaefer who served most capably as a research assistant.

I also wish to thank Virginia Grim for her excellent typing of the final draft of the manuscript.

On a more personal note, I wish finally to thank my family

for their unceasing help and encouragement.

October 1973

Daniel P. Keating
Baltimore, Maryland

TABLE OF CONTENTS

	Page
ABSTRACT	
PREFACE	
CHAPTER ONE	
Background and Statement of the Problem	1
CHAPTER TWO	
Method of Investigation	15
CHAPTER THREE	
Results and Analyses	32
CHAPTER FOUR	
Discussion: Conclusions and Implications	60
APPENDIX	
Letter to Parents	71
BIBLIOGRAPHY	73
FOOTNOTES	81
VITA	83

CHAPTER ONE

BACKGROUND AND STATEMENT OF THE PROBLEM

Theorizing about the nature of intelligence has a long and varied history, but theories which have been based on systematic observation of some sort are relatively recent. Two traditions currently exert a major influence on research and speculation about intelligence within psychology: the psychometric tradition and the Piagetian or Genevan tradition. They are clearly differentiated in their orientation, the theories they generate, and the methods which they employ.

Psychometrics

As its name correctly suggests, the psychometric tradition is based on the concept and practice of mental measurement. It is the topic of frequent reviews (e.g., Tuddenham, 1962; Bouchard, 1968; Robb, Bernardoni, and Johnson, 1972), perhaps the most cogent and complete of which is still Goodenough (1949). No single theory comprises the psychometric tradition. Instead, a number of theories which conflict to a greater or lesser degree lie within this tradition. The most noted are those of Spearman (1927), Burt (1949), and Vernon (1950) in Great Britain and Thorndike (1926), Thurstone (1938), and Guilford (1959) in the United States.

Characteristic of the psychometric tradition is the development and use of tests of "mental abilities." The tests consist of a collection of items which, it is hoped, tap a representative sample of the "universe of behaviors" to which the trait or ability to be measured gives rise. The "ability" may be quite general (e.g., abstract reasoning) or rather specific (e.g., clerical speed and accuracy).

The psychometric tradition is a direct descendant of the study of individual differences, as the above discussion would indicate. The union of the two is personified in Sir Francis Galton (1869, 1889). This dual interest in the differences in and inheritance of abilities predated many contemporary issues. The desire to measure reliably these abilities sparked the rapid growth of mental testing. The concern of such testing is what the individual is capable of producing in a structured and standardized situation, and the comparison of that production with the performances of other people in the same situation.

Typically the theories within the psychometric tradition are divided into two broad classes, one emphasizing general intelligence or "g," proceeding from Spearman to Burt to Vernon, the other emphasizing specific abilities, espoused by Thurstone and then Guilford. There are several things which all these theorists share in common, however, among them a belief that the test scores are maaningful numbers which can be manipulated mathemati-

cally to gain psychological insights, and also the underlying Darwinian assumption that variability exists in all traits and that the probability is high that the distribution representing this variability is essentially normal (Elkind, 1970).

In summary, then, the basis of the psychometric tradition is the measurement of individual differences in mental abilities or traits by means of the evaluation of a representative sample of behavioral products in a standardized situation. It has given rise to a diversity of theories.

Piaget

The Piagetian or Genevan tradition stands in contrast to the psychometric tradition on each of these points, although conceptual similarities in other areas do exist (Elkind, 1970). Jean Piaget, in work now spanning half a century, has offered a single, fairly unified theory of intelligence, although some of the details have of course been revised with time. The most concise outline of his theory is in the Psychology of intelligence (1950).

The primary goal of Piaget's investigations has been to identify those aspects of human thought which are universally characteristic of man. His approach has been at least as much biological and epistemological as psychological, but it is unified in that it is consistently structural and developmental. Although these various facets of his work are interesting in

their own right and deserving of study, it is principally the developmental aspect of his work which is of immediate concern.

Piaget (1950, etc.) has argued that the course of the individual's cognitive development involves different structures of thought at different stages. These stages, although corresponding roughly to age, are not defined chronologically. It is instead the different cognitive structures which define the different stages. Its four major periods of cognitive development are sensory-motor, pre-operational, concrete operational, and formal operational. The transition from the third to the fourth period is the concern of this investigation.

The determination of whether an individual's cognitive structures are those of one or another period is clearly of major importance to verification of the theory. Piaget has defined the concrete operational period and the formal operational period in terms of the logical structures which are available to the individual in them. As its name suggests, the concrete operational child can perform mental operations with respect to concrete (although not necessarily present) objects. The mental operations include the "classic" reasons the child gives upon achieving conservation: identity, negation, reversibility, and compensation. In general, concrete thought may be characterized as the appreciation of the unity of objects in reality despite changes; i.e., transformations of concrete objects no longer "fool" the child

(Inhelder and Piaget, 1958). The distinction between this and formal thought, at which period the individual presumably has available complete formal logic, is expressed by Inhelder and Piaget (1958):

> In sum, concrete thought remains essentially attached to empirical reality . . . . This characteristic of concrete thought is particularly clear in the present research . . . . The adolescent and the preadolescent [i.e., those beginning formal operations] start by formulating sets of hypotheses and then chooses experimentally the ones which are compatible with the facts. But, strictly speaking, the child does not formulate any hypotheses. He begins by acting; . . . he structures only the reality on which he acts [p. 250].

In order to assess the operational level of an individual for their research purposes, the Genevans have devised a series of cognitive tasks. By evaluating an individual's performance on these tasks according to the above criteria (e.g., formulation of hypotheses, reversibility of transformations), the individual's level of cognitive development is ascertained. A clinical approach is employed (methode clinique; Inhelder and Piaget, 1958, 1964; cf. Tuddenham, 1970; Mermelston, 1967) which involves a face-to-face interaction of the experimenter and the subject. It is not so much the correctness of the answer which is crucial, although it is taken into consideration, but the method of reasoning by which the subject came to his conclusion.

It should be clear that what is sought in such an approach

is not an adequate sample of a "universe of intelligence behaviors" but rather behavioral signs of an hypothesized underlying cognitive structure. A similar distinction was made by Goodenough (1949). An evaluation of the respective merits and demerits of the former (psychometric) and the latter (Piagetian) approaches to assessment is a second topic of the investigation.

In summary, Piaget's is a unified theory whose goal is the identification of universal characteristics of human thought. A major aspect of this approach is developmental, and the methodology is primarily clinical in that it seeks behavioral signs of underlying cognitive processes.

Precocity

One of the most consistently replicated findings of psychology is that on any dimension used to observe systematically the behavior of individuals there will be individual differences. It is not surprising then that in the area of cognitive development distinct individual differences are observed in the rate of such development. Regardless of the crudeness or sophistication of the observation, some individuals are clearly classed as "slow" developers and others as "fast" developers.

This is, in fact, the basis of mental measurement of intellectual ability within the psychometric tradition. Binet's breakthrough was to appreciate the usefulness for global intellectual

evaluation of these observably different rates of development (cf. Terman, 1916; Goodenough, 1949). Although Binet's original purpose was much more specific, namely the identification of those students too slow to benefit from regular school instruction, his perception of the relationship between precocity or retardation and eventual (i.e., adolescent or adult) intellectual attainment has been consistently verified empirically. The understanding of the relationship between precocity and brightness has not gone much beyond this empirical verification, however.

Since his concern has not been with individual differences, the role which precocity or retardation in mental development plays in Piaget's theory is relatively minor. It has been suggested by Piaget (1950) that an accurate evaluation in terms of developmental level could play an important part in an understanding of mental retardation.

It should be noted that the rate of cognitive development is not the only manner in which individual differences in intelligence can be conceptualized. As Inhelder (1953) noted, besides the rate of development, one could consider the extensivity of the cognitive structures throughout the various realms of human thought, the applicability of the structures to practical situations, or the flexibility which the individual has in using the structures available to him in novel ways. It is the utility of the rate measure which has led to a lack of interest in other

ways of viewing individual differences in intelligence. One attempt has been the research seeking to identify aspects of cognitive styles (e.g., Kagan, 1965) and their relationship to "intelligent behavior."

Pinard and Laurendeau (1964), in describing their attempt to develop a usable intelligence scale on the principles in Piaget's theory, suggest that

> The . . . problem of mental precocity itself appears in a new light if one can utilize an analytical instrument sufficiently nuanced and flexible to facilitate the development of accelerated scholastic programs without risking the upsetting of harmonious and integrated development [p. 260].

The Study of Mathematically and Scientifically Precocious Youth at Johns Hopkins has concerned itself with this latter problem, the identification and evaluation of mathematically precocious students and the development of "accelerated scholastic programs." Through the use of specific (rather than global) measures of cognitive abilities, a number of students has been so identified and the results in accelerating their educational programs have been encouraging (Keating and Stanley, 1972; Stanley, Keating, and Fox, in press). The assessment techniques used all fall within the rubric of "psychometric testing" (Keating, 1973), and much of the practical success of the project has been due to its treatment of precocious achievement in mathematics as a fait accompli. The tantalizing developmental implications of these

students' cognitive abilities have not been investigated directly, however, and that is a third topic of the current investigation.

Although precocity in intellectual development has been noted more widely in psychometric testing than in Piagetian investigations, this may be an artifact of the wider use of psychometric tests rather than an underlying difference in the results obtained using the two methods. In fact, in probably his most recent public statement on the topic, Piaget (1972) noted that there is great variance among individuals in the ages at which they acquire the different cognitive structures. It would seem that, at least in this case, when the same question is addressed by the psychometricians and the Piagetians the seemingly vast differences between them prove to be less formidable.

With respect to the topic of precocity in mental development, then, two questions need to be asked, one theoretical and the other practical. Students who by their age would be expected to be at the transitional stage between concrete operations and formal operations in Piaget's theory can be evaluated on psychometric tests and classed as either intellectually bright or intellectually average. The specific items by which they are so classed, however, may all be at the level of concrete operations. The items can require reasoning ability without requiring any formal operational reasoning. The question to be asked is whether intellectual brightness psychometrically defined implies cogni-

tive developmental precocity within a Piagetian framework.

The practical question involves the feasibility of developing a scaled series of tasks in order to evaluate an individual's level of intellectual development. At least two major attempts have been made to extend Piagetian principles in such a fashion (Pinard and Laurendeau, 1964; Tuddenham, 1970), but these and other less ambitious projects have concentrated on younger children (Pinard, personal communication). The development of tasks for assessing formal operations ability has been ignored by comparison, even by the Genevans, and the extent to which even Inhelder and Piaget's (1958) evidence supports their conclusions has been questioned (Bynum, Thomas, and Weitz, 1972). As Pinard and Laurendeau (1964) suggested, the uses of such a scale would be many, but few have investigated the possibility of doing so at the formal operational level.

A short digression regarding the importance of studying this period of development seems in order. The relative lack of concern with adolescent intellectual development stems from several sources, two of which deserve mention. The first is the assumption that early childhood experiences effectively "lock in" future behavior and that subsequent development flows more or less unchanged from them. Thus there is, e.g., the concentration of nearly all resources for remediation in this early childhood period. Rohwer (1971) has questioned the exalted status of this

assumption and argues that it is not at all clear that the course of intellectual development is irrevocably set in early childhood.

The second reason is perhaps even a less flattering reflection on psychological and educational researchers. When dealing with preschoolers or even elementary school children, the class of logical operations under investigation is relatively simple. In other words, the experimenter can without much mental strain be (almost) always right. When investigations of formal operations are undertaken the situation is quite different. Philosophers disagree as to whether or not a given operation should be considered an example of the exhibition of formal logic (e.g., Brainerd, 1971; Weitz, 1971; Bynum, Thomas, and Weitz, 1972).

There has, however, been a recent upsurge of interest in this adolescent period, sparked largely by Piagetian research. Recent studies have been reported by Bart (1971), Brainerd (1971), Furth and Youniss (1971), Higgins-Trenk and Gaite (1971), Jones (1972), Keasey (1971), Lunzer (1965), Martorano (1973), and others. These have dealt with the development of formal operations in one way or another, but little normative data is available from such studies because of the small N's and the diversity of tasks across studies.

At least two studies, one by Keasey (1971) and the other by Neimark and Slotnick (1970), have reported a curious finding which apparently contradicts an assumption stated by Inhelder and

Piaget (1958). The assumption is that all normal adults have available to them the operations of formal logic, although Inhelder and Piaget do not suggest that it is always used. The two studies mentioned did not always find evidence of formal operational reasoning in the adults they used as subjects. Most studies have used college students as their adult subjects, but Keasey (1971) used middle-aged women. (Her study used exclusively female subjects.) Formal operations were not inevitably observed in this group. A related speculation is that there is no reason to assume, as does Piaget (Inhelder and Piaget, 1958; Piaget and Inhelder, 1969) that formal operations mark the final stage of cognitive development. Riegel (1973) argues that another period of development, dialectic operations, needs to be assumed in order to complete Piaget's theory.

Few studies of which this investigator is aware have dealt with individual differences at age of acquisition of formal operations (Stephens, McLaughlin, and Mahaney, 1971; Saarni, 1972; Reiling, 1973; and Lovell and Shields, 1967). The small number of subjects and the varying criteria in these studies make any general conclusions from these studies highly tenuous. The individual differences in the Saarni study were field dependence-independence. Lovell and Shields studied gifted youngsters but reported no control group.

Thus it seems clear that there is little consensus with re-

gard to this period of mental development except perhaps among the Genevans. Consequently, great caution is required in undertaking and interpreting the results of Piagetian investigations at this period.

The Present Investigation

The overall goal of the present investigation is an examination of certain aspects of cognitive development at the period of preadolescence. This age has been selected because it is presumably the transitional period to formal operations in Piaget's theory (Inhelder and Piaget, 1958). Both psychometric and Piagetian techniques will be used in this investigation, and thus a comparison of their relative strengths and weaknesses for various purposes will also be undertaken. Finally, by preselecting the groups on cognitive variables psychometrically assessed, some further indication of the relationship between brightness and precocity, or at least a clarification of the questions involved, may result. From the above considerations, the hypotheses below were investigated:

1. The results of the psychometric and Piagetian assessments are essentially the same in terms of group differences. The intellectually "bright" and "average" groups, as determined by psychometric testing, are discriminable as such on the basis of the Piagetian assessment as well.

2. The reliability of the psychometric assessments for these groups exceed that of the Piagetian assessment for these groups, making the psychometric approach advisable in practical situations.

3. The precocity factor is as strong in the Piagetian assessment as in the psychometric assessment. Again the high groups demonstrate formal operational reasoning ability beyond their normal age group as determined from the evaluation of the "average" groups. In other words, "brightness" psychometrically defined implies precocity in the Piagetian developmental framework.

CHAPTER TWO

METHOD OF INVESTIGATION

In order to test empirically the hypotheses set forward at the end of Chapter One, several steps were essential. The first was to locate groups of early adolescents whose scores on psychometric tests were readily available. From this group it was necessary to select subgroups at different ages and at different levels of brightness psychometrically defined. After further screening on psychometric tasks, a reliable estimate of the cognitive developmental stage within Piaget's framework had to be made. These steps, which provided the data for the desired comparisons and evaluations, were carried out in the following manner.

Subjects

For reasons of proximity and previous contact, the Baltimore County School System was selected as being most likely to provide information on available subjects. The two grades which have students at about the right age to offer a test of the developmental hypothesis at hand were the fifth and the seventh grades. The students in these grades, tested in the summer, average about 11 and 13 years old, respectively, the ages specified by the

Genevans as the period in which the transition to formal thought occurs.

An additional benefit in using these students is that they had recently (in April) taken the Iowa Tests of Basic Skills (ITBS). Thus the psychometric information on which the initial screening was to take place would be relatively new.

This particular investigation was conducted with boys only. The use of male subjects exclusively stemmed from practical considerations but was justifiable on other grounds. The practical consideration involved the experimental design demands on the size of the group to be assessed. The necessary N would be at least doubled by the inclusion of female subjects. In addition, a female experimenter would be required for the female subjects, and two additional factors would have been introduced: sex of subject and sex of experimenter. This decision is justified by precedent (Keasey, 1971, and others) and by the empirical observation that extreme mathematical precocity at this age has been observed only in boys (Keating and Stanley, 1972; Astin, in press).

The arithmetic score of the ITBS was used for the initial screening. The arithmetic score was chosen, first, becuase of the apparently greater precocity observed in mathematical as opposed to verbal reasoning (Keating, in press) and, second, because it was thought to have a greater reasoning component than the verbal test, which in these grades tends to be oriented to vocabulary

rather than verbal *reasoning*.

The initial screening on the basis of the boys' ITBS arithmetic scores was in terms of the percentile ranks of Baltimore County norms. The "bright" groups were defined as 98th percentile or higher on those norms on the appropriate score; the "average" groups were defined as having scores between the 45th and 55th percentiles on county norms. The county norms are somewhat more stringent than the national norms, which meant that the "average" groups were slightly above average on national norms. If anything, this worked against the hypothesis of difference. National norms were not used because: 1) they were not easily available from the county schools; and 2) the national norm average groups would be below average for the county, possibly confounding a finding of difference with factors of low self-perception due to possible past experiences with school failure.

A print-out of all county students meeting the above criteria were received from the Baltimore County testing office.[1] The four groups at this point were: fifth grade bright; fifth grade average; seventh grade bright; and seventh grade average. A random selection procedure was used to choose fifty students in each of the four groups. These 200 students (50 in each of the four groups) were then invited to participate in a study of "science learning and evaluation techniques" for which they would be paid a small fee and also receive the score on the test they were to

take, and after which they might be recalled for an individual interview session for which they would again be paid. (A copy of the letter which went to the parents of these children is in the Appendix.) The random selection and invitation procedure was carried out by individuals other than the experimenter so that he could remain blind to the groups to which the subjects belonged.[2] The importance of this is described below (p. 27).

Of the 200 students who were invited, 109 showed up on the date of the testing. Several others asked for and were given make-up dates, but the conditions in these sessions were somewhat different from the large group testing, and these students were not included in the analyses which follow. The percentage of students in each group who did come on the testing date are shown in Table 1. As one can easily see, the younger students were somewhat more likely to come than the older ones, and the bright ones were much more likely to come than the average ones.

The differential ratios of "shows" and "no-shows" was expected, as in any study involving voluntary participation. Self-selection variables can affect the outcome of such investigations, but for two reasons it seems implausible that they did in this case. The first is that if a bias exists it is in favor of the brighter students in the average groups to come, thus closing the distance between them and the bright students. This, if it actually did occur, worked for the null hypothesis of no difference;

TABLE 1

Percent of Each of Four Groups Showing Up for Group Testing Session

	Average	Bright	Total
5th Grade	44%	74%	59%
7th Grade	38%	62%	50%
Total	41%	68%	

as seen below, a statistically significant difference was observed. The second reason is that the double screening procedure used increased the probability that the students who eventually participated did belong in the psychometric classifications to which they were assigned.

Following the group testing session, to be described below, an equal number of students was selected from each of the four groups for the Piagetian testing. The second screening used different criteria from the first. The purpose was to obtain groups which were clearly bright and clearly average on psychometric tests, and also to get groups which were reasonably homogeneous on age. Several students were eliminated from each of the four groups for one or the other of these reasons. The mean age in months for each of the four groups is shown in Table 2, separately for all who came to the group testing and for those who were selected for the follow-up testing.

Of the fifth graders who were retested, the average group (5A) had a mean age of 11 years, four months, and the bright group (5B) a mean age of 11 years, two months. This is the only noteworthy age difference in Table 2, but even this difference is not statistically significant. Note that in any event it worked against the hypothesis of precocity, since the average group was older. The mean age of both the seventh grade average group (7A) and the seventh grade bright group (7B) was 13 years, one month.

TABLE 2

Mean Age in Months of Students in Each of Four Groups, Unselected and Selected for Follow-Up Testing

	N	Mean	S.D.
Unselected 5A[1]	22	137.09	6.31
Selected	13	136.15	4.60
Unselected 5B	37	133.08	5.18
Selected	13	133.46	6.31
Unselected 7A	19	158.11	4.57
Selected	13	156.92	3.12
Unselected 7B	31	155.97	3.98
Selected	13	156.62	4.54

[1] 5A = Fifth grade average.

5B = Fifth grade bright.

7A = Seventh grade average.

7B = Seventh grade bright.

The psychometric test scores of the selected versus the unselected groups, which were additional criteria for selection, are discussed in the results section (p. 32).

Measures and Procedure

Since this investigation was designed in part to compare two different approaches to the same problem (i.e., human cognitive development), it was essential that adequate ways of measuring outcomes be used. Given this importance of the measures, a relatively detailed description of both the psychometric and the Piagetian measures is necessary.

Psychometric

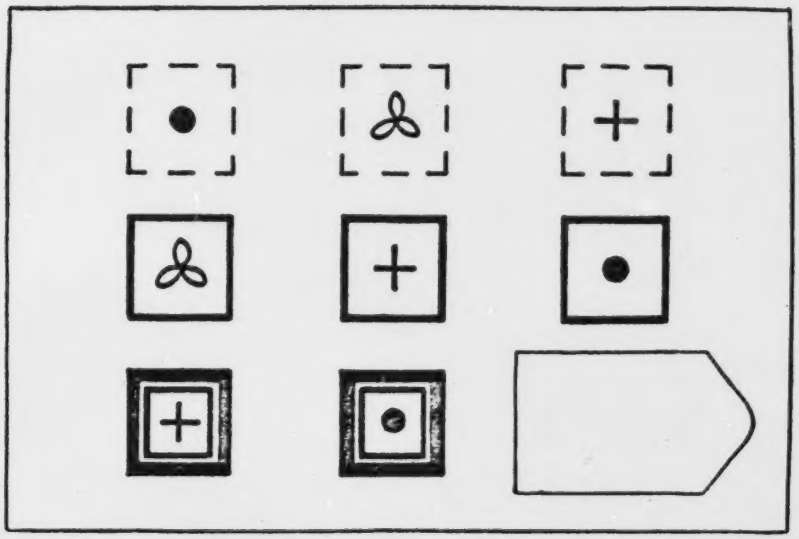
As mentioned above, the psychometric test which was used for the initial screening was the arithmetic score of the Iowa Test of Basic Skills. The arithmetic score is one of 15 subscores (Buros, 1965). The ITBS was designed as a measure of general academic skills. The test was administered to all fifth and seventh graders in the Baltimore County School System during April of 1973 in the regular county-wide administration. Baltimore County supplied the names of all the students in those grades who scored between the 45th and 55th percentile and those who scored above the 98th percentile of county norms.

The second psychometric assessment instrument which was used was Raven's Standard Progressive Matrices (Sets A, B, C, D, and E --Raven, 1938, 1956, 1960). This test was designed to be a "purer" measure of general intellectual ability (or "g," since Raven was an adherent to Spearman's two-factor theory) than the heavily verbally-loaded intelligence tests then in use, especially the Stanford-Binet (Raven, 1938). In fact, Spearman (1938) felt that the goal of measuring "g" had been achieved by this matrices test.

Regardless of one's theoretical orientation, the Standard Progressive Matrices, SPM, appears to be an excellent test of abstract reasoning ability. Court (1972) has collected a bibliography of over 300 SPM studies. Perhaps more remarkably, Piaget has stated (Inhelder and Piaget, 1964) that the "outstanding contribution to the field of multiplicative classification is, of course, Raven's 'progressive matrices' [p. 281]." Items D6 and E9 are shown here as Figures 1 and 2. The first item is of approximately medium difficulty for 11 year olds, with 64 percent of the 5A group doing it correctly. The second item (Figure 2) is difficult even for the 7B group with only 42 percent of that group doing it correctly.

The SPM was administered on a Saturday morning and afternoon to the 109 students who showed up of the 200 who were invited to participate. The students had previously been randomly assigned

D6



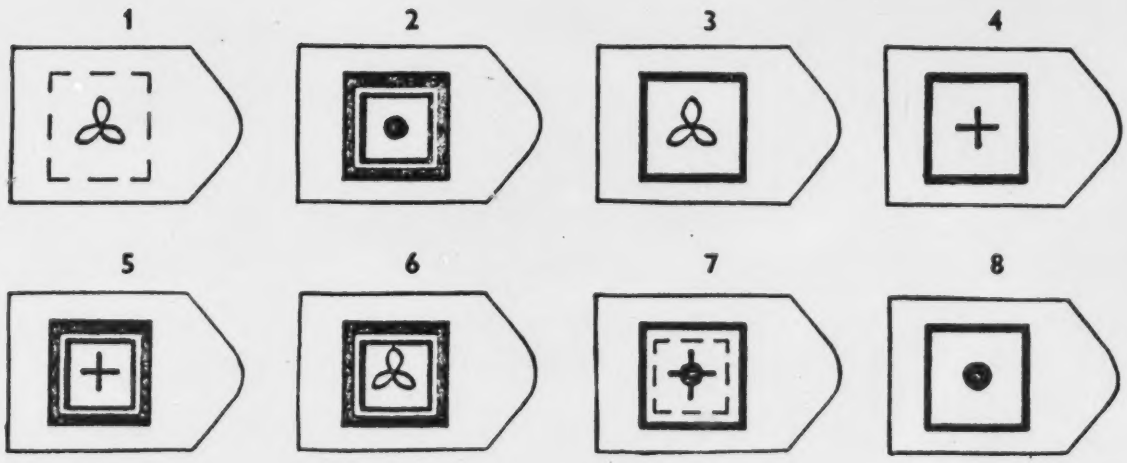
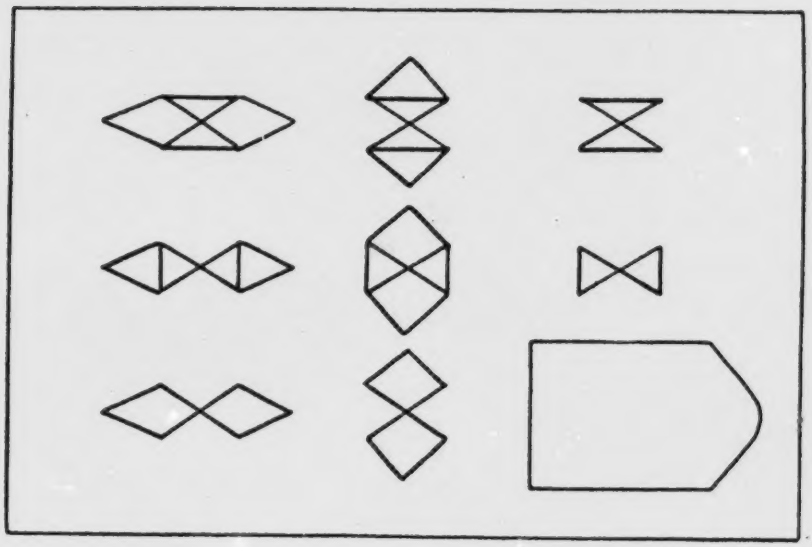


Figure 1: Item D6 from the Standard Progressive Matrices.

E9



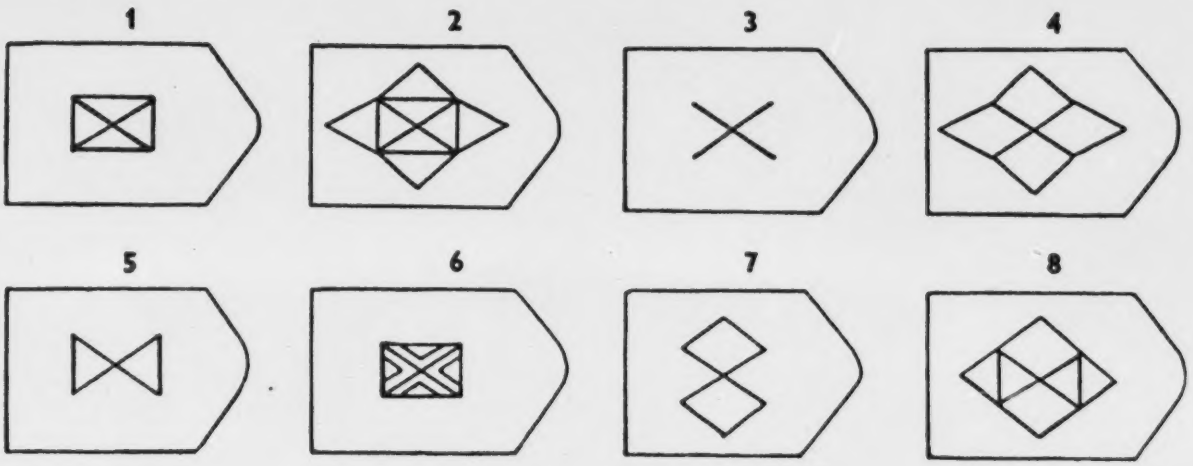


Figure 2: Item E9 from the Standard Progressive Matrices.

to either the morning or the afternoon session. The standard group administration procedure as stated in the SPM manual was followed without alteration.

It is also worth noting that "multiplicative classification" is not considered a formal operational level task (Inhelder and Piaget, 1964). Thus it is theoretically possible to do all the SPM items without using formal operations.

Both of these measures, the ITBS and the SPM, lie clearly within the psychometric tradition. On the basis of the description set forth in Chapter One (p. 3), they were both designed specifically for the measurement of individual differences through the evaluation of a (hopefully) representative sample of behavioral products in a standardized situation. Both are generally recognized as sound psychometric instruments (Buros, 1965).

Piagetian

In contrast to the psychometric measures discussed above, the Piagetian assessment presented several difficult problems of administration and scoring. The first was the selection of tasks in which formal operational thought, if present, could be used and observed. Inhelder and Piaget (1958) describe 15 tasks which they used in assessing formal operations. Although it has been questioned (Bynum, Thomas, and Weitz, 1972) whether or not the research described there conclusively demonstrates the presence

of formal thought, it was necessary in this investigation to use those tasks for which Piagetian protocols were available. One advanced concrete operational task (conservation of volume) and three formal operational tasks were used, and they are described later.

Consultation with a methode clinique-trained tester alleviated many problems involving detailed procedure.[3] Bias in scoring was avoided since the experimenter and rater were blind as to the group in which each subject belonged.

Another problem was that of subjectivity in the scoring of performance on the Piagetian assessment tasks. The scoring method used was adapted from Inhelder and Piaget (1958): II-A, concrete operational, a score of 0; II-B, advanced concrete operational, a score of 1; II-B-III-A, beyond concrete operational but not truly transitional to formal operational, a score of 2; III-A, transitional to formal operational, a score of 3; and II-B, formal operational, a score of 4.

Both the experimenter and the rater, who was present in the room at the time of the follow-up Piagetian interview, scored the protocol upon the completion of each task to be described below.[4] In the case of disagreements, both listened to the audio tape which was recorded during each session (with the knowledge and consent of the subject!) and arrived at a consensus score. The scoring criteria were also adapted from Inhelder and Piaget (1958).

The four Piagetian tasks that were used are described below:

1. Conservation of volume, which is an advanced concrete operational task and yields II-B as its highest score. This task is needed in order to determine whether or not the student has yet reached equilibrium in the concrete operational stage.

Two clay balls were formed by each child until he was satisfied that they were the same. Two containers of water were also adjusted until the S agreed they were the same. He was then asked to specify whether the water level would be the same or different after the clay balls were placed in the water, one in each jar. If S said "the same," it was demonstrated to him that he was correct. Then the clay balls were removed and one was made into a different shape ("sausage" or small pieces). The S was again asked whether the water levels would still be the same when the two clay balls were placed in the water. If he said "the same" and maintained the conviction upon questioning, he was scored as II-B. Otherwise, he was scored II-A.

2. Oscillation of a pendulum, in which the variables introduced were length of the string and weight of the object. As Inhelder and Piaget (1958) describe the task:

> The variables which, on seeing the apparatus, one might think to be relevant are: the length of the string, the weight of the object fastened to the string, the height of the dropping point (= amplitude of the oscil-

> lation), and the force of the push given by the subject. Since only the first of these factors is actually relevant, the problem is to isolate it from the other three and to exclude them. Only in this way can the subject explain and vary the frequency of oscillation and solve the problem [pp. 68-69].

3. Equilibrium in the balance. Again, in Inhelder and Piaget (1958) there is a description of this task:

> In a problem using a simple balance-type weighing instrument, a seesaw balance, we again find the operational schema of equilibrium between action and reaction. But the experiment was set up in a way that would force the question of proportionality. When two unequal weights W and W' are balanced at unequal distances from an axis L and L' , the amounts of work WH and W'H' needed to move them to heights H and H' corresponding to these distances are equal. Thus, we have the double (inverse) proportion:
>
> $W/W' = L'/L = H'/H$
>
> The result is that finding the law presupposes the construction of the proportion $W/W' = L'/L$ and spelling out its explanation implies an understanding of the proportion $W/W' = H'/H$ [p. 164].

4. Displacement. The purpose of this task was to see whether or not the child had the ability to separate variables such as weight, volume, and type of material enough to arrive at the concept of displacement. According to Inhelder and Piaget (1958),

> . . . the simplest contradictions can be overcome by means of concrete operations alone, the elimination of the more subtle ones, and particularly the formulation of a

> unified explanation, requires the use of implications--i.e., the intervention of formal propositional operations [p. 21].

The order of tasks was the same for each subject: advanced concrete task of conservation of volume; formal operational tasks of displacement, balance, and pendulum. There were not enough subjects to permit randomization of order. There are $3! = 6$ possible orders, which would mean a maximum of two subjects per order, far too few for suitable analysis. There is, however, no *a priori* reason to expect interaction of tasks and order. None have been reported in the literature.

Restatement of Hypotheses

Translating the hypotheses which were mentioned at the end of Chapter One in accord with the way they were investigated, the specific hypotheses were that:

1. There would be a clear difference in abstract reasoning between the groups as measured psychometrically by the SPM. The order of the groups was predicted to be: 7B > 7A; 7B > 5B; 7A > 5A; and 5B > 5A. No specific prediction was made about the relative standing of the 7A and 5B groups.

2. The groups would be ordered in an identical or very similar fashion on the Piagetian measures of cognitive developmental level, thus reaffirming the brightness-precocity relationship.

3. The reliability and efficiency of the SPM would surpass those of the Piagetian measures.

CHAPTER THREE

RESULTS AND ANALYSES

The results are presented here in three sections: first the analysis of the psychometric testing, the Raven's SPM data; second the analysis of the Piagetian clinical interviews concerning formal operational reasoning; and finally comparisons between them.

Psychometric

Order of the Groups

The means of the four groups (7B, 7A, 5B, 5A) on the SPM are shown in Table 3. They are in the order as predicted at the end of Chapter Two, and all the differences are statistically significant with the exception of the 7B > 5B difference. The relationship of the 7A and 5B groups was not predicted, but it is clear from Table 4, which lists the predicted order of the means, the actual comparisons, the obtained mean differences, and the levels of significance, that the 5B group is superior to the 7A group on SPM. It is also clear from Table 3 that the difference is more than 1 s.d.

The significance levels of the obtained differences should be considered approximately correct despite the fact that multiple comparisons were made, for two reasons. The first reason is

TABLE 3

Means of Four Groups on Raven's Standard Progressive Matrices, with Standard Deviation and Size of Group

		Average	Bright
7th	N	19	31
	$\bar{X}$	43.58	48.42
Grade	S.D.	4.81	4.25
5th	N	22	37
	$\bar{X}$	37.77	47.73
Grade	S.D.	6.58	4.53

Note: Highest possible score on SPM = 60.

TABLE 4

Predictions, Actual Comparisons, and Obtained Mean Differences of Four Groups on Standard Progressive Matrices Scores

Prediction	Actual Comparison	Mean Difference	Significance
7B > 5B	7B > 5B	.69	NS
7B > 7A	7B > 7A	4.84	$p > .001$
7B > 5A	7B > 5A	10.65	$p > .001$
None	5B > 7A	4.15	$.02 > p > .05$[1]
5B > 5A	5B > 5A	9.96	$p > .001$
7A > 5A	7A > 5A	5.21	$.01 > p > .02$

[1]This value was checked using Scheffé's multiple comparison method.

that the comparisons to be made were stated a priori and their direction predicted; thus a normal (two-tailed) t-test is a conservative test of difference. Second, the p values are quite small in nearly every case. The exception to these restrictions is the 7A, 5B comparison, in which direction was not predicted and for which therefore the significance level is not as clear. Accordingly, Scheffé's multiple comparison method (in McNemar, 1962) was used in examining this difference, and the significance level remained unchanged.

If then we were to write out the order of the groups on the SPM results, the order would be

$$7B = 5B > 7A > 5A .$$

That is, the 7B and 5B groups are equal on the SPM results; both are significantly higher than the 7A group; and all three are significantly higher than the 5A group. This relationship is depicted graphically in Figure 3.

The failure to obtain a significant difference between the 7B and 5B groups is surprising. The means of the two groups are not near the ceiling of the test, which has 60 items, but each individual in the group may be near his "personal ceiling." Neither does regression toward the mean following a 98th percentile cutoff criterion seem to explain this result adequately, since regression should affect the two groups equally. The result may simply reflect the different self-selection variables within the

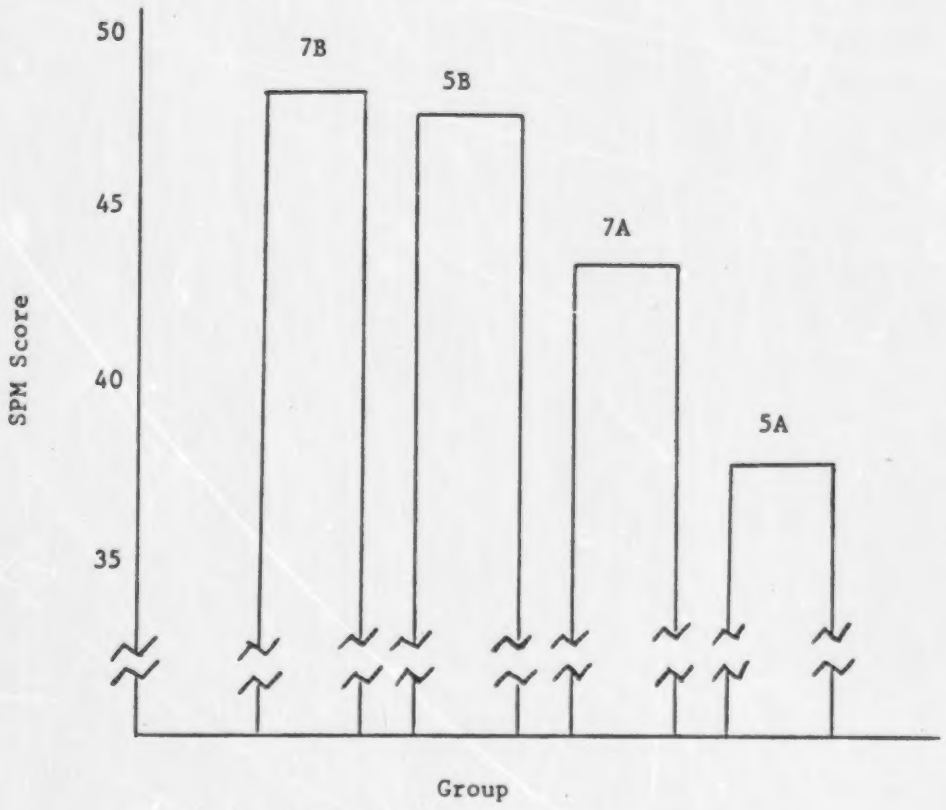


Figure 3: Scores of four groups on SPM.

two groups, or it may be that a score in the 50's on SPM requires much more intellectual development than a score in the high 40's.

Reliability

The SPM is composed of five different sets or scales: A, B, C, D, and E. The reliability of each group was determined from the intraclass correlation coefficient of the scales stepped up via the Spearman-Brown formula. The intraclass coefficient of correlation of the scales, which can be taken as the "typical" scale in the test (Stanley, 1971, p. 398), is

$$\rho_i = \frac{\overline{\rho_{ij}\sigma_i\sigma_j}}{\overline{\sigma_i\sigma_i}} .$$

A lower bound of the one-form reliability coefficient is secured by stepping up $\hat{\rho}_i$ by the formula

$$\frac{S\hat{\rho}_i}{1 + (S - 1)\rho_i} ,$$

where S equals the number of scales of the test (in this case, 5)[5] and $\hat{\rho}_i$ is the sample estimator of ρ_i. In Table 5 are listed the SPM reliability coefficients for each of the four groups, together with the value of $\hat{\rho}_i$.

Since the reliability is obviously based on the variance and

TABLE 5

Intraclass Coefficients of Correlation and Internal-Consistency Reliability Coefficients for Four Groups on SPM.

	5A	5B	7A	7B	All Groups[1]
Intraclass $\hat{\rho}_i$	.40	.17	.21	.21	.25
Reliability	.78	.50	.56	.57	.63

[1]The calculation of the reliability coefficient for "all groups" is made by stepping up the mean $\hat{\rho}_i$, rather than by taking an arithmetic averaging of the reliabilities for each of the four groups, which is inappropriate.

covariance within any single group, the prior restriction of range in the selection procedure tends to depress the reliability estimates. The value reported in the SPM manual (Raven, 1956) for an unselected group of 12-14 year olds is .88. None of the values in the four groups are that high, and the average reliability for the four groups is .63. This relatively lower reliability of course does not attenuate the main purpose of the test, which was not to separate individuals within the groups reliably, but to separate the groups themselves on the basis of reasoning ability psychometrically assessed. It should also be noted that the much higher reliability of the 5A group may be a sampling fluctuation, an artifact of the data, rather than a psychologically meaningful difference.

The discrepancy is, however, clearly related to the greater variance on SPM scores of the 5A group. The variance for the four groups is not significantly heterogeneous, using Hartley's test for homogeneity of variance. This somewhat greater variance of 5A may account for the "personal ceiling" on SPM, since it is possible that individuals in the other groups may score as high as they are ever likely to, while some 5A subjects may still be developing the "multiplicative classification" ability which SPM tests.

Piagetian

Inter-Rater Agreement

Before the data for the Piagetian assessment can be analyzed with any degree of confidence, it is necessary to ascertain whether or not the data are meaningful in terms of the behavior which occurred in the clinical interview. This can be done to some extent in terms of the degree of inter-rater agreement. If the two individuals scoring the subjects (i.e., the experimenter and the rater) in terms of the criteria discussed in Chapter Two agree consistently on the score each subject received, then confidence in at least this aspect of the data is justified.

The percentage of agreement of the two raters is listed in Table 6. There was exact agreement in the vast majority of the cases (130/156 = 83%), and disagreement by more than one point in either direction in only one of the 156 ratings. The correlation coefficient between the scores of the two raters was +.94. It differed slightly for the four groups, ranging from +.91 to +.96. Thus it seems safe to conclude that the scoring criteria were sufficiently specific and objective to permit a high level of agreement between independent scorers. Disagreements were resolved by consensus after both raters listened to the audio tapes of the sessions in question.

Order of the Groups

TABLE 6

Percent Agreement of Two Scorers on 156 Ratings of Reasoning Level.

	Exact Agreement	Differ by 1	Differ by 2	Differ by 3	Total
N	130	25	1	0	156
%	83.3	16.0	.7	0	100

Consider first the results of the conservation of volume task. None of the students who were tested was scored concrete as opposed to advanced concrete. (These ratings, incidentally, did not figure into the inter-rater agreement calculations.) This is somewhat surprising for the 5A group, but it should be remembered that the A groups are actually somewhat above the national average for fifth graders.

The means of the four groups on total score for the three formal operational tasks are shown in Table 7. The number in each group was 13, a total of 52 altogether.[6] The predicted and actual order of the groups is shown in Table 8 along with the level of significance for each comparison. Clearly the 7B and 5B groups are equivalent in terms of total score on the formal operational assessment. It is less clear whether they should be considered superior to the 7A group, since the more conservative two-tailed t-test yields only a trend in both cases ($.05 < p < .10$) while a one-tailed t-test, which is reasonable since an a priori prediction was made, yields a significant difference ($p < .05$). All three groups are in turn superior to the 5A group on formal operations total score. If we were to write out the relationship of the groups on this dimension, it would read

$$7B = 5B \geq 7A > 5A .$$

The developmental precocity of the bright versus the average groups is depicted graphically in Figures 4 and 5. Within each

TABLE 7

Means of Four Groups on Formal Operations Total Score, with Standard Deviations

		Average	Bright
Seventh Grade	$\bar{X}$	6.62	9.46
	S.D.	2.87	3.13
Fifth Grade	$\bar{X}$	4.46	9.15
	S.D.	1.39	2.48

Notes: Highest possible score for Formal Operations = 12.

For each group, $n = 13$; total group, $N = 52$.

TABLE 8

Predictions, Actual Comparisons, and Obtained Mean Differences of Four Groups on Total Formal Operations Score.

Prediction	Actual Comparison	Mean Difference	Significance
7B > 5B	7B > 5B	.31	NS
7B > 7A	7B > 7A	2.84	.05 > p > .10
7B > 5A	7B > 5A	5.00	p > .001
5B > 7A[1]	5B > 7A	2.53	.05 > p > .10
5B > 5A	5B > 5A	4.69	p > .001
7A > 5A	7A > 5A	2.16	.02 > p > .05

[1]Derived from Table 4, recalling hypothesis (on p. 30) that order of groups on Piagetian testing would duplicate that of psychometric testing.

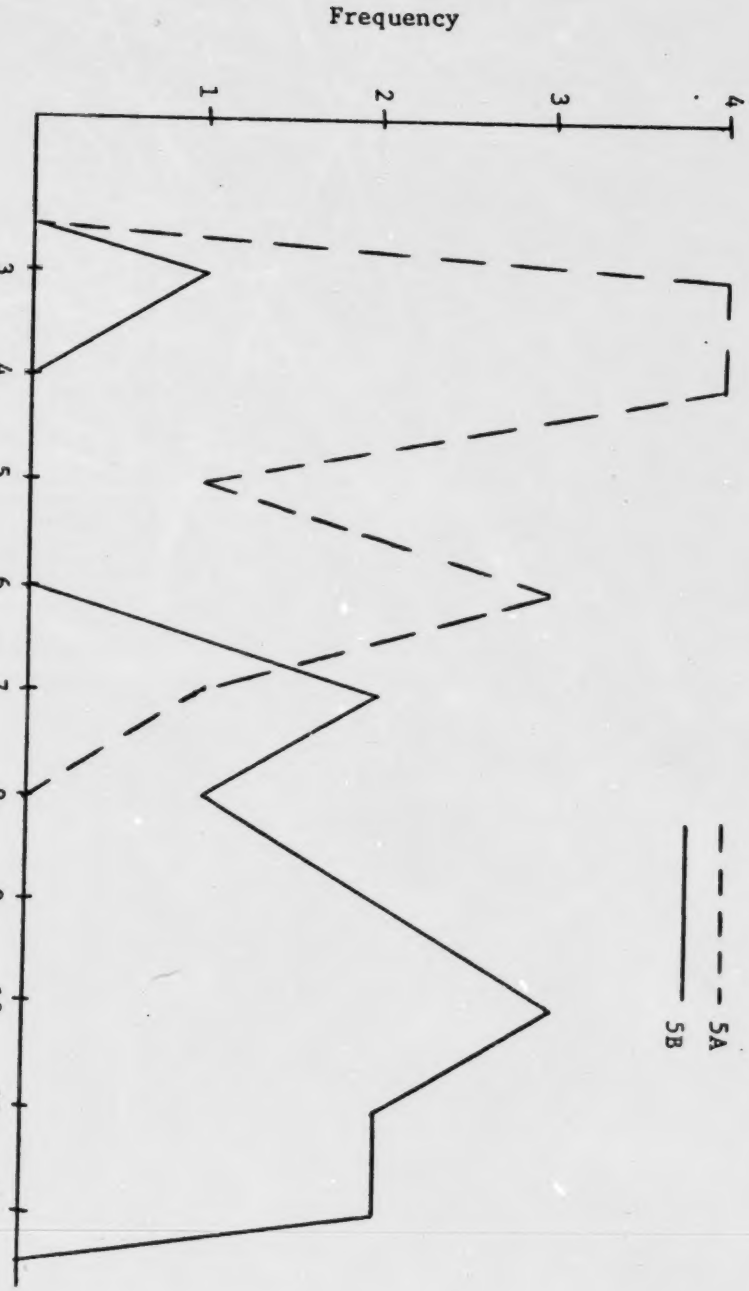


Figure 4: Scores of bright and average fifth graders on formal operations.

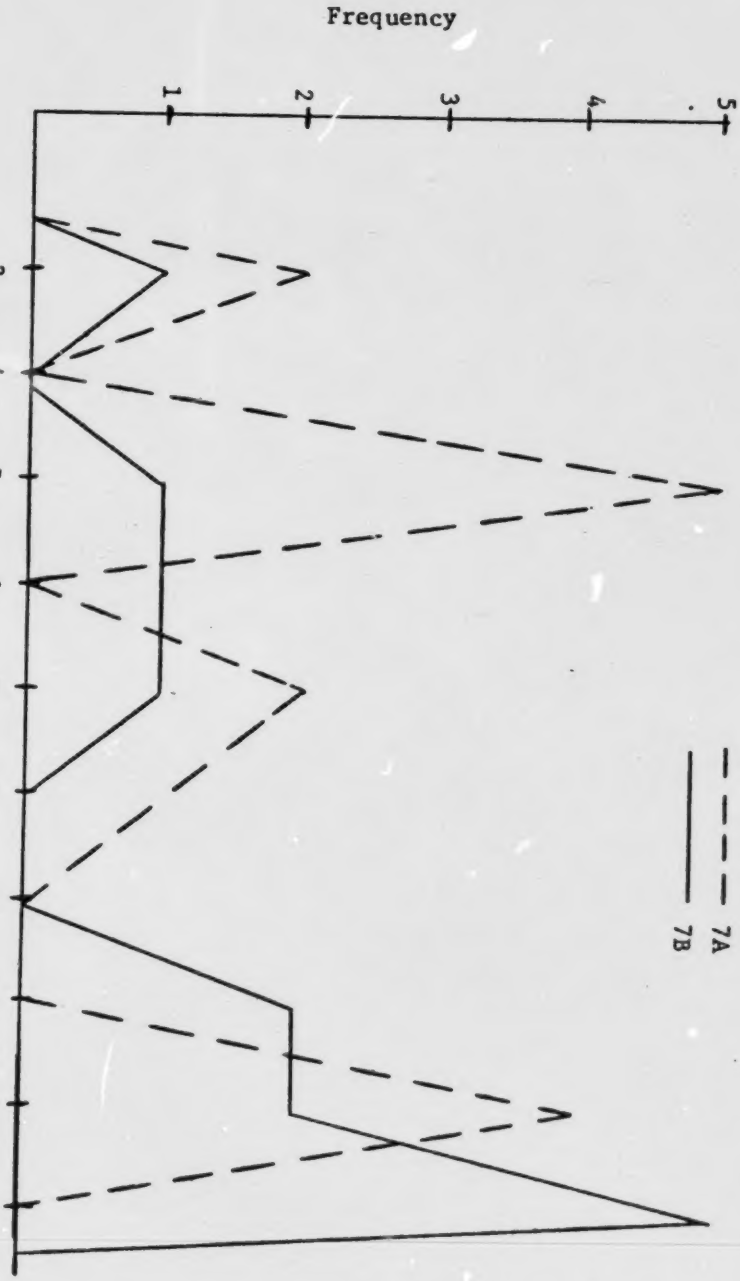


Figure 5: Scores of bright and average seventh graders on formal operations.

grade the frequency of the total scores is plotted. None of the 5A group scored above the median of the 5B group, and only one of the 5B subjects (7%) scored below the median of the 5A group. In the seventh grade the results were not as clear-cut, with two in the 7A group (15%) above the 7B median, and four in the 7B group (30%) below the 7A median. This is somewhat to be expected since a number of average 13 year olds (represented by the 7A group) do seem to use formal operations (Inhelder and Piaget, 1958).

The concept of a total score for formal operations assessment is anathema to many Piagetian researchers. They would prefer to compare the groups in terms of percent achieving formal operations. In Table 9 are listed the percent of each of the four groups who demonstrated formal operations on each of the three tasks, and also the percent showing formal operations on any of the three tasks. The overall result from Table 9 in words is that the 7B and 5B groups have attained formal operations, while the 7A group is transitional and the 5A group is still advanced concrete operational.

The relationship of the four groups indicated by the data in Table 9 is

$$7B = 5B > 7A > 5A .$$

The reliability of these differences is not entirely clear, but the significance of the differences in the total score was discussed above. In terms of previous Piagetian research, a criter-

TABLE 9

Percent of Each of Four Groups Demonstrating Formal Operational Reasoning on Each of Three Tasks[1]

	5A	5B	7A	7B
Displacement	0	70	38	77
Balance	8	85	47	77
Pendulum	23	70	38	62
Total[2]	31	93	62	85

[1] This is defined as a score of 3 or 4 on the task.

[2] Not a sum of the above three. It is the percent of different students in each group showing formal operations on <u>any</u> of the three tasks.

ion of 75 percent of the group is common for determination of the cognitive developmental level of a group. By this criterion, the developmental precocity of the bright groups is apparent: 7B and 5B are both above the criterion, and 7A and 5A are both below it. The percentages from Table 9 are shown for comparison in Figure 6.

Note from Table 9 and Figure 6 that on the basis of the percentages there the 5B group actually appears to be superior to the 7B group. This is, however, somewhat misleading. Two of the 7B subjects did not seem to get involved with the problems at all, thus lowering the overall percentages for the 7B group.

A different way of looking at the same data is to consider the number of students within each subgroup who scored at a formal operational level (i.e., a 3 or a 4) on all three of the formal operational tasks. These percents for each of the four groups are: 5A, 0%; 5B, 46%; 7A, 23%; and 7B, 62%. Thus in this analysis the 7B group appears superior to the 5B group, in that they have extended these formal operations to more difficult areas.

These various analyses of the same data point up one of the major pitfalls in the interpretation of results such as these. Without a detailed and consistent criterion from one study to another, there seems little hope of assimilating findings from different studies. Yet another way of examining the data is with a more conventional psychological tool, the analysis of

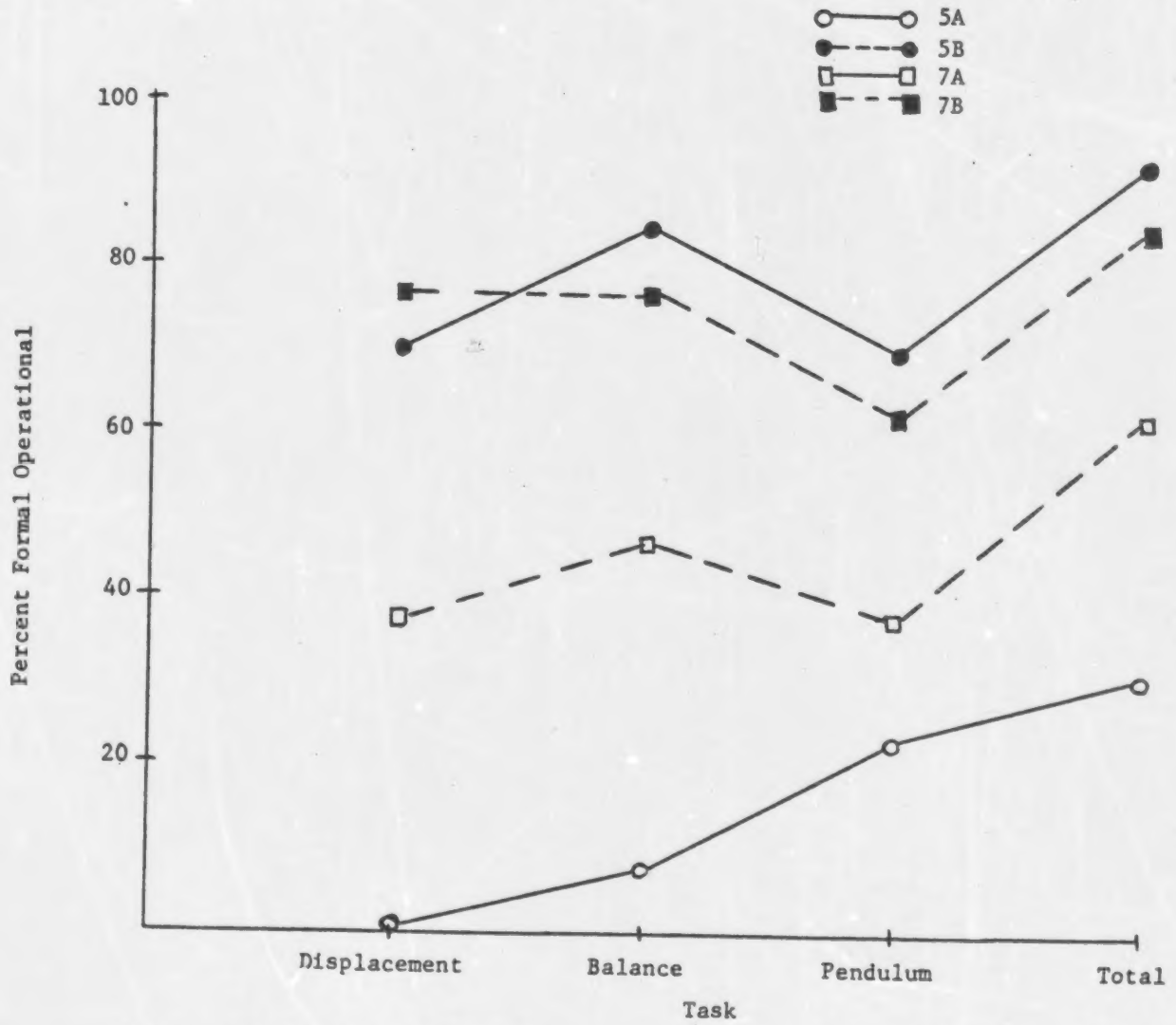


Figure 6: Percent in each of four groups demonstrating formal operations on each of three tasks.

variance.

The effect of the psychometric selection of the groups is seen clearly in an analysis of variance (ANOVA) of the scores on the Piagetian tasks contained in Table 10. The variable of psychometric brightness is highly significant ($p < .001$), but the variable of age fails to reach significance by a slight margin ($.05 < p < .10$). This latter is undoubtedly due to the transitional nature of the 7A group, which reduces the distance between them and the 7B group, and also to the superiority of the 5B group over the 7A group. This would seem to indicate that within an age group which is transitional from one stage to the next, i.e., 11 to 13 year olds, intelligence (in the psychometric sense) is far more important than age in determining the cognitive developmental level of the individual.

There was, somewhat surprisingly, no main effect for tasks. None of the tasks was more difficult than the others across all groups. One might suspect a "halo effect" in the scoring were it not for two other pieces of evidence: first, the high percentage of inter-rater agreement; and second that fully 25 percent of the subjects received a strictly concrete operational score on one

TABLE 10

ANOVA of Scores on Formal Operational Tasks

Source of Variance	df	SS	MS	F
Psychometric level (P)	1	61.603	61.603	30.119**
Age-Grade (G)	1	6.603	6.603	3.234*
PG	1	3.653	3.653	1.794
S/PG	48	97.744	2.036	
Tasks (T)	2	2.186	1.093	1.465
PT	2	2.070	1.035	1.387
GT	2	2.993	1.497	2.007
PGT	2	.443	.222	.298
ST/PG	96	71.641	.746	

*$.05 < p < .10$

**$p < .001$

task (a score of one) and a transitional or formal operational score (three or four) on another task.

The ANOVA of the formal operations scores produced no significant interactions.

Reliability

The three formal operational tasks were each considered to be one item for the purpose of obtaining the reliability of the Piagetian assessment. As in the SPM reliability, the Piagetian assessment reliability was computed separately for each group. The formula is that for the stepped-up intra-class coefficient of correlation as noted above for the SMF (p. 37), but with items rather than scales.

Before examining the actual figures, a cautionary note should be added. This method of estimating reliability was developed for psychometric testing, and is based on its theory and practice. The homegeneity of the items is not an assumption for successful assessment within the Piagetian tradition, however. Piaget has used the term "horizontal decalage" to describe the observed fact that individuals do not always or even usually extend their newly acquired operational constructs to all possible situations at once. Within the concrete operational stage, for example, there appears to be a clear progression of conservational ability through the different dimensions of length, volume, area and so

forth.

This provides three options for examining what psychometricians call reliability. The one more in conformance with Piagetian theory would be to see whether the tasks under consideration are scalable and maintain that scale when used on different groups. That possibility is ruled out in this case, since there was no consistency in the difficulty of the tasks across groups. The second possibility is of course a test-retest design to look at stability of the measures. This must be ruled out for two reasons: first, if too long a period of time intervenes between test and retest, the possibility of differential developmental maturation of the subjects is severe; and second, if the retest is done shortly after the initial test, there is the problem of memory of the tasks.

This leaves only the third option of examining the homogeneity of the items. In Table 11 are listed the intraclass coefficients of correlation and the reliability of the assessment for each of the four groups. As one can see from Table 11, the values vary a great deal across groups. The particular pattern is difficult to understand. The unreliability of the measurement for the 5A group is probably due to the lack of any true formal operational reasoning in the group. In those few cases where individuals obtained scores greater than one, it was essentially random. The extremely high values for the 7B group appear to be

TABLE 11

Intraclass Coefficients of Correlation
and Reliability Coefficients for
Four Groups on Test of Formal Operations

	5A	5B	7A	7B	All Groups[1]
Intraclass $\hat{\rho}_i$	-.06	.29	.43	.80	.37[2]
Reliability	-.08	.55	.69	.92	.63

[1]The calculation of the reliability coefficient is made by stepping up the mean $\hat{\rho}_i$, rather than by taking an arithmetic average of the reliabilities of the four groups, which is inappropriate.

[2]This value was checked using the ANOVA formula for estimating $\hat{\rho}_i$, which is (Stanley, 1971, p. 400):

$$\hat{\rho}_i = \frac{MS_p - MS_{pxi}}{MS_p + (I - 1)MS_{pxi}},$$

where MS_p is the mean square for persons, MS_{pxi} is the mean square for persons by items, and I equals the total number of items (in this case, 3).

an artifact of several unexpectedly low scores in that group for two individuals, thus increasing the homogeneity of the items. The 5B and 7A figures appear to be more representative of the reliability one might obtain in an unselected group with a reasonable amount of variance on the aspect of cognitive development measured by these tasks.

Comparisons

There are traditionally a number of ways of examining data such as these, and frequently different interpretations can be made from the same data set. In this investigation, however, the major result regardless of the method of analysis is quite clear: brightness psychometrically defined implies cognitive developmental precocity within Piagetian theory and practice.

As mentioned above, the order of the groups on performance in both the psychometric and the Piagetian tasks is as predicted, with the only difference between the two in the statistical significance of the mean differences. This order can be written for the SPM results as

$$7B = 5B > 7A > 5A ,$$

and for the Piagetian assessment results as

$$7B = 5B \geq 7A > 5A .$$

In terms of overall <u>efficiency</u> of measurement, the SPM was clearly superior, which is not surprising since it was developed

within a measurement-oriented framework. When one considers that two or three individuals can test 100 or more children in a two-hour period effectively and reliably enough with the SPM to separate bright and average groups, and that the same man-hours are necessary to achieve the same goal for one child with the Piagetian measures, then the difference is quite obvious. It is interesting to note, however, that the reliability of the two types of measures was the same, .63. The direct comparison is biased against the Piagetian measures, since the number in each group was far fewer than for the SPM. The reliability of the SPM for just those individuals would be theoretically a better comparison were it not for the fact that the SPM was itself used in the selection, thus reducing its variance drastically.

Perhaps the fairest comparison of the two measures is in their relative effectiveness in separating the four groups from each other, as opposed to the reliability which indicates how well the individuals within a group are distinguished from each other. The latter is adversely affected by the homogeneity of the group, whereas the former is positively affected by the same factor. In this case the psychometric measurement is more effective (see Tables 3 and 8). But this is again somewhat confounded by the relatively lower "ceiling" of the Piagetian test.

Although more detailed investigation of this specific point would seem warranted, these results clearly suggest that the Pia-

getian measures are not inherently poorer even for the detection of individual differences than the psychometric measures designed for that purpose, but that the time required for their administration makes them impractical. The Piagetian measures do have the signal advantage, however, of being tied to a unified, coherent, theoretical position, whereas the psychometric measures are not.

The relationship of the variables of age and brightness (as represented by the SPM score) with cognitive developmental level can be examined directly through their correlations with each other. The correlations within groups of age and SPM score with formal operations total score are listed in Table 12. None of the correlations is significantly different from zero correlation, which is not surprising given the restricted nature of the groups.[7] But in the average groups some trends are apparent on both the age and the intelligence variables. In an unselected population of the appropriate age range, these correlations would likely be significant.

Suffice it to say that for these data no clear-cut relationship in terms of individual scores between the SPM and performance on the formal operations tasks exists which is not confounded with group effects. Note however that in the clearly transitional group, 7A, the correlations are in the expected direction although not statistically significant.

TABLE 12

Correlation of Formal Operations Score
with Age and SPM Score in Each of Four Groups

	5A	5B	7A	7B
Age	.48*	-.30	.42	.16
SPM	.05	.10	.55*	.11

*$.05 < p < .10$

CHAPTER FOUR

DISCUSSION: CONCLUSIONS AND IMPLICATIONS

The major hypothesis of this investigation was that brightness as measured by psychometric testing implies developmental precocity in reasoning. This precocious reasoning was to be confirmed by comparisons of bright and average groups both on further psychometric testing and, more importantly, on measures of cognitive developmental level which have been used by Piagetian researchers. This hypothesis was confirmed as the results presented in Chapter Three indicated.

This confirmation extended to the unexpected result regarding the 5B and 7A groups. Even though they averaged two full years younger, the 5B students were at a more advanced cognitive developmental level than the 7A students, as witnessed by their superior performance on both the psychometric test of abstract reasoning ability, the SPM, and on the Piagetian analysis of cognitive development.

The results regarding the secondary hypothesis of greater reliability and efficiency of the psychometric measure were not as clear. The psychometric measure was more consistently reliable across groups, but this is at least partially the result of adequate variance in all the groups on SPM, whereas the 5A group had decidedly less variance than the other groups on the Piage-

tian assessment. The average reliability of the four groups on the SPM and the Piagetian assessment was about the same. In terms of efficiency, however, the psychometric measure was clearly superior, as is to be expected.

This should not be interpreted as reflecting negatively on the Piagetian assessment, however, since it was designed for a different purpose from that of the psychometric measures. Piagetian assessment possesses the distinct advantage of being closely tied to a coherent theory within which the results from such assessments are clearly interpretable, which is not always the case with psychometric measurement.

As stated earlier, the major purpose of the investigation was to examine the cognitive developmental level of early adolescents who are at a transitional stage within Piaget's framework, with the addition of an individual difference variable. The technical superiority of measures derived within a measurement oriented theory was not the main concern, but it is not surprising to find them to be superior.

If one were to analyze in the other direction, so to speak, by doing a Piagetian analysis of items from psychometric measures, the theoretical superiority of the Piagetian system would perhaps be just as apparent. The items on the Stanford-Binet Intelligence Scale, for example, show dramatic shifts of item content at the ages during which Piaget suggests there are stage changes in

cognitive development. Webb (1973) has suggested that this is perhaps the major reason for the characteristic "wash-out" of IQ gains obtained during special preschool programs, a difficult phenomenon to explain from a puremeasurement approach. This type of training is typically for the items which correspond to the preoperational stage in Piaget's framework, and when the shift to "concrete operational" items occurs at ages six or seven, then the "gains" are lost. Thus an empirically derived scale shows these shifts, incorporated originally because of their effectiveness in showing individual differences, which are accounted for by an external theory.

If one judges theories by what they do (e.g., in terms of measuring their important concepts) rather than by what they say, the relationship between rival theories is on occasion made less obscure. The manner in which the theories under consideration in this research developed led to wide differences in the way in which they spoke of "intelligence." But many of the differences are more apparent than real when a direct investigation of the comparison is carried out. This is not to deny that there are differences in the aspects of human intelligence which they concentrate on, or even that these differences are important for a complete understanding of intelligence, but only to suggest that there is not as great a difference in practice, in what people really do in the real world, as one might suspect from what the

theories say about the subject.

Implications for Practice

There are at least two important practical implications which result from this study, and both are concerned with the relationship of psychometrics and Piagetian theory.

The first implication is that the results of this investigation offer little encouragement to proceed with the development of a "Piagetian scale of intelligence" analogous to the psychometric measures now in existence, at least at the level of formal operations. Although the Piagetian tasks used in this study reveal no inherent unreliability of measurement, there seems to be little purpose in designing a new scale which, in terms of individual differences in intellectual functioning, will reveal little more than current instruments. A more worthwhile goal would seem to be careful analysis of existing psychometric instruments from several different theoretical points of view in order to understand their inner workings on a firmer basis.

For example, there seems little necessity to develop a "Piagetian scale of intelligence" in order to be able to identify individuals who are precocious in cognitive development. High scores for one's age group on psychometric measures are adequate measures of such precocity, especially if several psychometric instruments are used to minimize errors of measurement (e.g.,

Stanley, Keating, and Fox, in press). Psychometric brightness does imply cognitive developmental precocity.

The second practical implication is that researchers using psychometric measures to discover precocity may rest assured that they do, in fact, accomplish that task. The strictly empirical nature of most psychometric tests does not change their efficiency in detecting individual differences in the pace of cognitive development.

Implications for Theory

Perhaps the single most surprising result in the Piagetian portion of the investigation is the failure to obtain a main effect of tasks. The experience of Piagetian researchers with the concrete operational stage has been quite different. Not only are there major differences in the difficulty levels of the various concrete operational tasks, but the order of difficulty is quite consistent across studies. This often replicated finding has been used as evidence for the universality of the stage theory, and of concrete operations in particular.

Since this has been the past experience with the immediately preceding stage, concrete operations, the absence of any order of difficulty raises important questions about the viability of the assertion of a structure-related stage change in early adolescence.

These questions must be raised for other reasons as well.

Not only is there an absence of an order of difficulty which is consistent across groups, but this coexists with the presence of what Piaget has called "horizontal decalage," which is the failure of individuals to extend the newly acquired structure immediately to all possible situations. If there were no horizontal decalage, one could argue that this is the reason for the absence of order. But such was not the case.

In addition to these considerations is the fact that previous research outside Geneva on formal operations has tended to raise more questions than it has answered, which is quite a different case from the research on concrete operations. Of particular note is the failure to obtain cross-cultural replication (e.g., Goodnow, 1969).

All these considerations, taken together with impressions gathered from observation of the students' performances on these tasks, lead to the suggestion that the undeniable change in intellectual functioning which takes place at the beginning of adolescence is not truly a change in the cognitive structure of the individual, but rather a maturation of cognitive strategies which corresponds to an overall psychological maturation taking place at this time. This would seem to account for the lack of consistency in the difficulty level of the tasks. If it were truly a <u>structural</u> change, as seems to be the case in concrete operations, one would expect a progression in order from one set

of tasks to the next. A strategy maturation, however, would permit for variance in how well the new strategy is applied in different situations.

Another piece of evidence against a structural stage change is that in at least one study of adult subjects, the presence of formal operations was not observed (Keasey, 1971). One could argue that the individuals did possess formal operational ability but did not demonstrate it in that particular experimental framework, but that seems a somewhat implausible ex post facto rationalization.

If this stage change at adolescence were one of strategy rather than of structure, it suggests several possibilities for instruction. It might be something which after an as yet undetermined point in development would be teachable. Thus perhaps adolescence is an equally valuable if not more valuable period for remediation when compared with early childhood. It certainly suggests further research to determine if the structure vs. strategy distinction is relevant to the issue at hand.

This investigation also implies that a synthesis of the two traditions, psychometric and Piagetian, is not impossible. Although only the outlines are even vaguely visible at this point, it is clear that there can be no complete understanding of human intelligence without taking into consideration both the universals of cognitive development, which Piaget has attempted to iso-

late and describe, and the undeniable fact of individual differences in human abilities.

These results do not yield much further information on the nature of the brightness-precocity relationship other than to reaffirm the empirical correlation across different traditions. This reaffirmation does provide the opportunity for some speculation on the point, however.

The most straightforward explanation would seem to be that brightness leads to precocity. If an individual possesses a greater amount of "g," to use Spearman's term, then his interaction with the environment is likely to produce more information of a high quality, since he manipulates it in more varied and interesting ways, and, further, it is more likely that he will be able to incorporate this new information more effectively and efficiently. The net result, of course, is that he will progress through the stages of cognitive development more quickly, since, in Piagetian terms, he will encounter contradictions sooner, thus bringing about disequilibrium within a stage or at the transition point to the next stage at an earlier time than the average. The same explanation would of course apply to the question of slow development: poorer information is generated by the interaction of the dull individual with his environment, and he is less able to make use of the information which is produced.

One final implication for theory should be noted. Goodenough

(1949) recounted the following analogy which generated so much interest among psychologists in the cognitive development of retarded children: the cognitive development of retarded children is of particular interest because it is like a slow moving train compared with the rapid pace with which normal development proceeds. Thus through a careful study of the intellect of retarded children they hoped to gain an insight into the normal development of the intellect. This hope has been largely abandoned with the recognition that the differences between normal and retarded development are often in kind rather than in degree.

It may be that the metaphor is more heuristically valuable if it were to be turned around. The pace of normal cognitive development may be too slow for us to discern the nature of the major changes which occur along the way to maturation. It may be that our most valuable insights into the nature of cognitive development will result from the careful study of bright children, where the processes may be more identifiable. This of course remains to be seen.

Future Research

The results of this investigation suggest a number of different areas in which further research is definitely indicated. Four of the possible research topics are discussed here.

The first necessity is of course for similar studies to be

carried out across a broader spectrum of individual differences, both qualitatively and quantitatively. A good example of the former is a study by Saarni (1972) which attempted to relate the development of formal operations to the dimension of field dependence or independence. In the latter instance, the quantitative dimension, i.e., the age of development of formal operations at several different levels of intelligence should be studied to get a better idea of the true relationship than can be gained from essentially a four point comparison such as this investigation.

The second is the necessity to study the teachability of formal operational concepts (or strategies). The research on this topic is extensive in the concrete operational stage, and the issue there is still controversial (Elkind, 1970). The preponderance of evidence at this point seems to indicate that concrete operational concepts are not teachable so that they remain with the individual for any length of time. But if the formal operational development is not structural, then it is possible that this difference from concrete operations would be revealed in instructional experiments. A good population with which to start (theoretically although probably not pragmatically) is non-formal operational adults. The second group would then be transitional adolescents of different ability levels, and finally concrete operational adolescents.

A third area of investigation is suggested by the above, and that is whether the structure-strategy distinction holds up under further scrutiny. At the moment there exists no widely accepted vocabulary with which to talk about individual differences in cognitive strategies. The term cognitive styles is purposely avoided here to avoid confusion with the dimensions already discussed by others under that rubric, and which are probably permanent characteristics of the individual closely related to personality variables.

The fourth area which bears further investigation is the nature and quality of the interaction of the individual with his environment, and how this is related to individual differences in intelligence. Some variables which may be useful in this regard have been developed by White (1972), who prefers to speak of overall competence rather than intelligence and whose major concern is with maternal-child interaction as an input variable. It is probable that such research must begin with natural or controlled natural observations to attempt to isolate the variables of interest. It does not preclude making these observations on older children and adolescents, however, since cognitive development presumably continues throughout this period.

APPENDIX

LETTER TO PARENTS

THE JOHNS HOPKINS UNIVERSITY

Baltimore, Maryland 21218

Department of Psychology

June 8, 1973

Parents of

Dear Parents:

Your son has been selected from a list of Baltimore County youngsters as a potential participant in a study of science learning and evaluation techniques. The one or two sessions for which we are requesting your assistance should be interesting, enjoyable, and helpful for your son, and most definitely not harmful in any way. Since these students were carefully selected from a much larger list, it is essential to the study that as many as possible who were selected participate.

The first part of the study will consist simply of a paper-and-pencil test about abstract designs, in which the student learns as he goes along. Most youngsters find it a very interesting and "different" test. Your son is scheduled for the

___ Morning session (8:45-11:30 a.m., Shaffer Hall, Room 3)

___ Afternoon session (1:00-3:45 p.m., Shaffer Hall, Room 3)

The date of the session is Saturday, July 14, 1973. The location of Shaffer Hall is shown on the enclosed map.

For giving up half a summer Saturday, your son will be reimbursed $2.00 as a token of our appreciation. Please indicate on the enclosed form whether or not your son will be coming on July 14th.

On the enclosed form you will also note the question concerning availability for a follow-up study. The follow-up will involve a single individual session of about two hours duration (maximum) with some "science learning" tasks, which should prove quite interesting for these students. About 30 percent of the students who come on July 14th will be invited to a follow-up session, which will be held from Monday through Friday, July 30 to August 17, 1973. Those students who are invited back and agree to participate will receive $4.00 for their cooperation. If your son is coming on July 14th, would you also indicate whether or not he would be available for a follow-up session from July 30 to August 17? (If he can come only on a weekend or evening between those dates, please check "Yes" and indicate weekend and/or evening at the bottom of the form.)

Participation in the study should prove rewarding and enjoyable for your son. We will of course send you information concerning results of the sessions. If you have any questions, please call me at 366-3300, Ext. 1410.

Sincerely yours,

Daniel P. Keating
Project Associate

DPK:vsg

Encl.: Reply form and envelope
Map of the campus

BIBLIOGRAPHY

Astin, H. S. Sex differences in mathematical and scientific precocity. In J. C. Stanley, D. P. Keating, and L. H. Fox (Eds.), Mathematical talent: Discovery, description, and development. Baltimore, Md.: Johns Hopkins Univ. Press, in press.

Bart, W. M. The factor structure of formal operations. British Journal of Educational Psychology, 1971, 41, 70-77.

Bouchard, T. J. [Jr.] Current conceptions of intelligence and their implications for assessment. In P. McReynolds (Ed.), Advances in psychological assessment. Palo Alto, Calif.: Science & Behavior Books, Inc., 1968.

Brainerd, C. J. The construction of the formal operations of implication-reasoning and proportionality in children and adolescents. Dissertation Abstracts, 1971 (May), Vol. 31 (11-B), 6919-6920.

Buros, O. K. The sixth mental measurements yearbook. Highland Park, N.J.: The Gryphon Press, 1965.

Burt, C. The structure of the mind: A review of the results of factor analysis. British Journal of Educational Psychology, 1949, 19, 110-111 and 176-199.

Bynum, T. W., Thomas, J. A., and Weitz, L. J. Truth-functional

logic in formal operational thinking: Inhelder and Piaget's evidence. Developmental Psychology, 1972, 7(2), 129-132.

Edwards, A. J. Individual mental testing. Scranton, Pa.: International Textbook, 1969.

Elkind, D. Children and adolescents: Interpretive essays on Jean Piaget. New York: Oxford Univ. Press, 1970.

Fox, L. H. Facilitating educational development of mathematically precocious youth. In J. C. Stanley, D. P. Keating, and L. H. Fox, (Eds.), Mathematical talent: Discovery, description, and development. Baltimore, Md.: Johns Hopkins Univ. Press, in press.

Furth, H. F. and Youniss, J. Formal operations and language: A comparison of deaf and hearing adolescents. International Journal of Psychology, 1971, 6(1), 49-64.

Galton, F. Hereditary genius. London: MacMillan & Co., 1869.

Galton, F. Natural Inheritance. New York: The MacMillan Company, 1889.

Goodenough, F. L. Mental testing: Its history, principles, and applications. New York: Rinehart & Co., 1949.

Goodnow, J. J. Cultural variations in cognitive skills. In D. R. Price-Williams (Ed.), Cross-cultural studies. Baltimore, Md.: Penguin Books, Inc., 1969.

Guilford, J. P. Three faces of intellect. American Psychologist, 1959, 14, 469-479.

Higgens-Trenk, A. and Gaite, A. J. H. Elusiveness of formal operational thought in adolescence. Proceedings of the 79th American Psychological Association, 1971, 201-202.

Inhelder, B. Criteria of the stages of mental development. In J. M. Tanner and B. Inhelder (Eds.), Discussions on child development. Vol. I. New York: International Universities Press, Inc., 1953.

Inhelder, B. and Piaget, J. The growth of logical thinking from childhood to adolescence. (Translated by A. Parsons and S. Milgram.) New York: Basic Books, 1958.

Inhelder, B. and Piaget, J. The early growth of logic in the child. London: Routledge and Kegan Paul, Ltd., 1964.

Jones, P. A. Formal operational reasoning and the use of tentative statements. Cognitive Psychology, 1972, 3, 467-471.

Kagan, J. Reflection-impulsivity and reading ability in primary grade children. Child Development, 1965, 36, 609-628.

Keasey, C. T. The nature of formal operations in preadolescence, adolescence, and middle age. Dissertation Abstracts, 1971 (June), Vol. 31 (12-B), 7574.

Keating, D. P. The study of mathematically precocious youth. In J. C. Stanley, D. P. Keating, and L. H. Fox, (Eds.), Mathematical talent: Discovery, description, and development. Baltimore, Md.: Johns Hopkins Univ. Press, in press.

Keating, D. P. Discovering quantitative precocity. Unpublished

paper, 1973.

Keating, D. P. and Stanley, J. C. Extreme measures for the exceptionally gifted in mathematics and science. Educational Researcher, 1972, 1(9), 3-7. (Sept.)

Lipsitt, L. P. and Eimas, P. D. Developmental psychology. In P. Mussen and M. Rosenzweig, (Eds.), Annual Review of Psychology, 1972, 23, 1-50.

Lovell, K. and Shields, J. B. Some aspects of a study of the gifted child. British Journal of Educational Psychology, 1967, 37, 201-208.

Lunzer, E. A. Problems of formal reasoning in test situations. In P. Mussen (Ed.), European research in cognitive development. Monographs of the Society for Research in Cognitive Development, 1965, 30(2), 19-46.

Martorano, S. C. The development of formal operational thinking. Paper presented at the biennial meeting of the Society for Research in Child Development, Philadelphia, March, 1973.

McNemar, Q. Psychological Statistics. New York: John Wiley & Sons, 1962.

Mermelston, E. A note on Piaget's clinical method of questioning. Alberta Journal of Educational Research, 1967, 13(3), 181-184.

Neimark, E. and Slotnick, N. S. Development of the understanding of logical connectives. Journal of Educational Psychol-

ogy, 1970, 61, 451-460.

Piaget, J. Creativity. Paper presented at the Milton S. Eisenhower Symposium, The Johns Hopkins University, Baltimore, October, 1972.

Piaget, J. Psychology of intelligence. London: Routledge and Kegan Paul, Ltd., 1950.

Piaget, J. and Inhelder, B. Intellectual operations and their development. In Oleron, P., Piaget, J., Inhelder, B., and Greco, P. (Eds.), Intelligence. Vol. VII of Experimental psychology: Its scope and method, P. Fraisse and J. Piaget, (Eds.), London: Routledge and Kegan Paul, Ltd., 1969.

Pinard, A. and Laurendeau, M. A scale of mental development based on the theory of Piaget: Description of a project. Journal of Research in Science Teaching, 1964, 2, 253-360.

Raven, J. C. Progressive Matrices (1938), Sets A, B, C, D, & E. London, H. K. Lewis & Co., Ltd., 1938.

Raven, J. C. The 1956 Revision of the Matrices Test. Bulletin of the British Psychological Society, 1957, 32, 3.

Raven, J. C. Guide to the Standard Progressive Matrices. London: H. K. Lewis & Co., Ltd., 1960.

Riegel, K. F. Dialectic operations: The final period of cognitive development. Research Bulletin, 1973, RB-73-3, Educational Testing Service.

Reiling, A. M. The effect of an individual difference dimension

on the development of formal thought. Paper presented at the biennial meeting of the Society for Research in Cognitive Development, Philadelphia, March, 1973.

Robb, G. P., Bernardoni, L. C., and Johnson, R. W. Assessment of individual mental ability. Scranton, Pa.: In text, 1972.

Rohwer, W. D. Prime time for education: Early childhood or adolescence? Harvard Educational Review, 1971, 41(3), 316-341.

Saarni, C. A. Piagetian operations and field independence as factors in children's problem solving performance. Paper presented at the annual meeting of the American Educational Research Association, Chicago, March, 1972.

Spearman, C. E. Measurement of intelligence. Scientia, 1938, 64, 75-82.

Stanley, J. C. Reliability. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Stanley, J. C. Intellectual precocity. In J. C. Stanley, D. P. Keating, and L. H. Fox, (Eds.), Mathematical talent: Discovery, description, and development. Baltimore, Md.: Johns Hopkins Univ. Press, in press.

Stanley, J. C., Keating, D. P., and Fox, L. H. (Eds.). Mathematical talent: Discovery, description, and development. Baltimore, Md.: Johns Hopkins Univ. Press, in press.

Stephens, B., McLaughlin, J. A., and Mahaney, E. J. Ages at which

Piagetian concepts are achieved. Proceedings of the 79th American Psychological Association, 1971, 203-204.

Terman, L. M. The measurement of intelligence. Boston: Houghton, 1916.

Thorndike, E. L. The measurement of intelligence. New York: Bureau of Publications, Teacher's College, Columbia, University, 1926.

Thurstone, L. L. Primary mental abilities. Chicago: Univ. of Chicago Press, 1938.

Tuddenham, R. D. Psychometricizing Piaget's methode clinique. In I. J. Athey and D. O. Rubadeau (Eds.), Educational implications of Piaget's theory. Waltham, Mass.: Ginn and Co., 1970.

Tuddenham, R. D. The nature and measurement of intelligence. In L. Postman (Ed.), Psychology in the making. New York: Alfred A. Knopf, 1962.

Vernon, P. E. The structure of human abilities. London: Methuen, 1950.

Webb, R. A. The evaluation of a style-oriented cognitive curriculum for two- and three-year olds in the I. V. Y. Program of the Baltimore City Public Schools. Unpublished report, 1973.

Weitz, L. J. A developmental and logical analysis of Piaget's sixteen binary combinations. Dissertation Abstracts, 1971

(Oct.), Vol. 32 (4-B), 2412.

White, B. L. An analysis of excellent early educational practices: Preliminary report. Interchange, 1971, 2(2), 71-88. (Cited in Psychological Abstracts, 1972, 48(2), 291.)

FOOTNOTES

[1]I wish to thank Dr. John H. McCauley in the Baltimore County School System Office of Testing for his cooperation in supplying the initial information required for the selection of groups. His assistance is gratefully acknowledged.

[2]I wish to thank William C. George for organizing the selection of groups and the contacting of the individual students.

[3]Grateful acknowledgement is made of several lengthy conversations with Dr. Francois Stoll.

[4]I wish to thank Rosalind A. Schaefer for serving as rater for the Piagetian sessions and assisting in the preliminary data analysis.

[5]The one-form reliability coefficient for the SPM would be exact if the intraclass coefficient of correlation of the items rather than the scales were used, but the difference between it and this estimate would be minimal.

[6]In three of the groups 14 individuals were assessed with the Piagetian measures, but it was not possible to get the 14th subject in the fourth group (5A). In order to maintain an equal n , one subject was randomly eliminated from each of the other three groups.

[7]In some instances it is possible to correct for restriction of range on the correlation coefficient, but in order to do so

some estimate of the variance of an unselected group is required. To my knowledge, no suitable estimates exist for these comparisons.

VITA

Daniel Patrick Keating

I. Personal

Date of Birth: January 10, 1949
Marital Status: Single
Social Security No.: 212-48-3436

II. Education

Elementary, Archdiocese of Baltimore Schools
Secondary, Loyola High School, Towson, Maryland
B.A., College of the Holy Cross (psychology), 1971, cum laude
M.A. The Johns Hopkins University (psychology), 1973.

III. Professional Experience

Project Associate, Study of Mathematically and Scientifically Precocious Youth (SMSPY), The Johns Hopkins University, 1971-present.

Teaching Assistant, 3 semesters, Department of Psychology, The Johns Hopkins University, 1971-72.

Instructor, Intersession course, Department of Psychology, The Johns Hopkins University, 1973.

Instructor, Spring semester, Department of Psychology, Towson State College, Towson, Md., 1973.

Editor, Newsletter of SMSPY, 1972

IV. Papers and Publications

Keating, D. P. and Stanley, J. C. From eighth grade to selective college in one jump: Two case studies in radical acceleration. Presented at annual meeting of American Educational Research Association (AERA), Chicago, 1972.

Keating, D. P. The study of mathematically precocious youth. Paper presented at the annual meeting of the

American Association for the Advancement of Science (AAAS), Washington, D.C., 1972.

Keating, D. P. and Stanley, J. C. Discovering quantitative precocity. Paper presented at the annual meeting of the AERA, New Orleans, 1973.

Keating, D. P. Creative potential of mathematically precocious youth. Paper presented at the annual meeting of the American Psychological Association, Montreal, Quebec, Canada, 1973.

Keating, D. P. Possible sampling bias in Genetic Studies of Genius. Educational and Psychological Measurement, in press.

Keating, D. P. and Stanley, J. C. Extreme measures for the exceptionally gifted in mathematics and science. Educational Researcher, 1972, 1(9), 3-5.

Stanley, J. C., Keating, D. P., and Fox, L. H. (Eds.). Mathematical talent: Discovery, description, and development. Vol. I of Studies in Intellectual Precocity. Baltimore, Md.: The Johns Hopkins University Press, in press.

Keating, D. P. The study of mathematically precocious youth. In J. C. Stanley, D. P. Keating, and L. H. Fox (Eds.), Mathematical talent: Discovery, description, and development. Chapter 2. Revised version of AAAS paper.

Weiss, D. S., Haier, R. J., and Keating, D. P. Personality characteristics of mathematically precocious boys. In J. C. Stanley, D. P. Keating, and L. H. Fox (Eds.), Mathematical talent: Discovery, description, and development. Chapter 7.

Keating, D. P., Wiegand, S. J., and Fox, L. H. Classroom behavior of mathematically precocious youth. In J. C. Stanley, D. P. Keating, and L. H. Fox (Eds.), Mathematical talent: Discovery, description, and development. Chapter 9.